A Mixed Method Approach for Evaluating and Improving the Design of Learning in Puzzle Games

Monica Visani Scozzi
University College London
London, UK
monica.visani@gmail.com

Ioanna Iacovides
The Open University
Milton Keynes, UK
jo.iacovides@open.ac.uk

Conor Linehan
University College Cork
Cork, Ireland
conor.linehan@ucc.ie

ABSTRACT
Despite the acknowledgment that learning is a necessary part of all gameplay, the area of Games User Research lacks an established evidence based method through which designers and researchers can understand, assess, and improve how commercial games teach players game-specific skills and information. In this paper, we propose a mixed method procedure that draws together both quantitative and experiential approaches to examine the extent to which players are supported in learning about the game world and mechanics. We demonstrate the method through presenting a case study of the game *Portal* involving 14 participants, who differed in terms of their gaming expertise. By comparing optimum solutions to puzzles against observed player performance, we illustrate how the method can indicate particular problems with how learning is structured within a game. We argue that the method can highlight where major breakdowns occur and yield design insights that can improve the player experience with puzzle games.

Author Keywords
Games; Learning Curves; Breakdowns; Player Experience, Evaluation Methods; Games User Research.

ACM Classification Keywords
H.5.3 Information interfaces and presentation (e.g., HCI): Miscellaneous ; K.8.0. General: Games.

INTRODUCTION
One of the primary concerns of Games User Research (GUR) is to develop processes through which to evaluate and improve the player experience. Many GUR methods have been developed, from using heuristic review to identify critical issues, [e.g. 22] to using physiological measures in play-testing [e.g. 30]. However, while learning (defined here as learning 'how to play and progress') is often referred to within GUR methods, there is no established procedure intended specifically for identifying problems with how player learning is structured within commercial games, and for offering evidence-based solutions to these problems. This paper outlines and demonstrates a method appropriate for doing just that.

Many researchers, designers and commentators have argued that learning is a necessary part of all gameplay, from the understanding of game controls and interfaces, to working out solutions for in-game puzzles and challenges [i.e. 14, 15, 23]. Indeed, without developing expertise and an appropriate understanding of the game world, players will not be able to experience deeper levels of involvement such as flow [40]. However, empirical studies of learning in games focus almost exclusively on measuring the effects of game-playing on educational outcomes [7], and much more rarely on how to support learning 'how to play and progress' through design and scaffolding.

Recent work has started to address this issue, such as Linehan and colleagues [25], who examined puzzle games from a structural perspective. Through doing so they developed a set of "learning curves", which illustrated how commercially successful games are designed to introduce different skills to players and provide them with space to practice those skills. However, missing from their analysis was an observation of how players reacted to those designed curves.

Another relevant strand of research has examined gameplay in terms of the *breakdowns* experienced during play [19, 20]. This research has established processes for identifying points at which problems occur in gameplay, and for identifying the different strategies players adopt in an attempt to overcome those problems (i.e. examining how players achieve *breakthroughs*). Individual strategies can be used to diagnose the type of problems experienced by players and point to specific types of solutions. While this research provides a rich picture of how involvement is influenced by learning during play, it could be argued that such an in-depth approach would be too time consuming to adopt in a typical game development context.

We aim to build upon previous work around learning in games by presenting a mixed method approach for evaluating the player experience in commercial games from a learning perspective. We demonstrate our approach through a case study of *Portal*, a puzzle game, where we carried out an in-depth observational study with 14 players of varying expertise. Recordings of gameplay were used to

CHI PLAY '17, October 15–18, 2017, Amsterdam, Netherlands

ACM ISBN 978-1-4503-4898-0/17/10…$15.00
https://doi.org/10.1145/3116595.3116628

compare optimal performance to observed performance, while think-aloud and interview data helped to provide further insights into player experiences. We discuss specific levels and mechanics that players struggled with, before concluding with a consideration of design implications and a discussion of how our approach could be adopted to evaluate and improve the design of learning in both commercial and educational video games.

RELATED WORK

In this section, we present a brief literature review, which argues that learning is fundamental to the experience of playing a game. We discuss existing research on how to design games that players learn from consistently, and on methods that focus on evaluating learning in games. In addition, we consider existing methods used for evaluating player experience, and argue that there is an opportunity to improve our practice around the design and evaluation of learning throughout gameplay.

Learning & gameplay

Researchers, designers and commentators have long recognized learning as a core component of the experience of playing games [9, 15, 20, 25]. Through play, players learn to master interface controls, game rules and mechanics. Additionally, there is broad agreement that enjoyment of games depends largely on how well the challenges presented to players match their level of skill [e.g. 1, 6, 40], where games that are too difficult or too easy can lead to negative experiences [1, 38]. Thus, it is important to ensure, through design, that players have developed the necessary skills and information to complete the challenges that the game presents. In other words, we need to carefully design how the game teaches players.

Designing for learning in games

We are only beginning to see research that teases out how to design games to ensure that players are effectively and efficiently equipped with the necessary skills and understanding to complete in-game tasks and challenges. For example, Linehan et al. [26] present a summary of features found useful for the design of learning in games. Games benefit from the inclusion of a hierarchy of short, medium and long-term goals [38]. Long, complex tasks should be broken down into short, simple components that are trained individually before being chained together [14, 15, 23]. Players should be required to demonstrate competence at simple tasks before advancing to more complex tasks. Games should provide immediate feedback in a context appropriate manner [27, 28, 29].

It has also been argued that the *pacing* of challenges is central to player's ability to learn game-specific skills [8, 9, 13]. Pacing refers to the trajectory of difficulty over time as players advance through the game. Similarly, it is generally agreed that games should present challenges that are matched to the player's individual skill level, and that playing games is fun only if a sufficient proportion of the game challenges are mastered by the player [e.g. 1, 6, 15, 39, 40].

The challenge of designing learning progression in games is related to that of tutorial design [see 3]. Some games, such as *Portal*, incorporate all instructions and rehearsal into the general narrative of game play. Others use explicit tutorial segments and instructional materials. Either way, the same basic questions exist – when should new skills be introduced, how much help is necessary, what strategies ensure that players learn and master skills? Andersen et al., [3] carried out a study with 45,000 participants, where they questioned the "added value" of tutorials to game engagement and progression. Through investigating eight different designs, they found that tutorials have clear benefits in more complex games, while being less useful in simpler games. Thus, the authors were able to demonstrate the importance of providing structure to the skills and information learned in a game (i.e., using tutorials versus not) and for learning to be carried out through game play – and that the effect is pronounced in games that are more complex. We argue, based on an understanding of instructional design, that these concepts generalise to all game design, rather than applying simply to tutorials.

Linehan et al. [25] suggest that pacing is managed particularly effectively in commercially successful and well-regarded puzzle games. By analysing the structure of successful games, they found that each of the main skills are introduced separately through a simple puzzle that requires little more than the basic performance of the new skill. After this introduction, a number of puzzles are presented in which the player is given the opportunity to practice that skill and to integrate it with previous ones. From the point at which a new skill is introduced, puzzles increase in complexity up until the point at which the next new skill is introduced. These are useful principles around which to design learning curves in games, and match very closely with guidelines for the design of special education programmes [26]. However, missing from their analysis was an observation of how players reacted to those designed curves. It would be useful to understand whether this structured approach genuinely facilitated a satisfying experience for players. The current study will look at that question more closely.

Evaluating learning in games

The majority of research that examines games in the context of learning focuses on the use of games as educational tools and the gamification of virtual learning environments. Within the context of education, there has been a more recent focus on examining not just learning outcomes but how well a game is able to facilitate the process of learning. Many of these involve data mining, learning analytics and modelling techniques. For instance, Harpstead et al. [16], present an analysis of an educational game, which involved knowledge component modelling and empirical learning curve analysis. By combining statistical modelling with player performance data (from game logs), the authors provided insights into problems with game design and developed suggestions for improving the game to ensure that students were learning what they were supposed to.

Other examples include Andersen at al. [2] who developed the Playtracer method, where multidimensional scaling is used to produce visualisations of how groups of players move through a game, and Owen et al [32] who present the ADAGE (Assessment Data Aggregator for Game Environments) protocol for tracking players interactions and the leveraged machine learning, data mining and statistical methods to assess learning within play.

Outside of formal education, there has also been a focus on investigating the process of learning within commercial games, where researchers have examined the problems that players encounter and how to overcome them [4, 24, 35]. For instance, Iacovides et al. [18, 20] build upon work by Sharples [37] to introduce the notion of gameplay breakdowns and their converse, breakthroughs. Breakdowns and breakthroughs are described as occurring in relation to player *action* (e.g. problems with the controller vs performing a new attack); *understanding* (e.g. not knowing what to do next vs figuring out a solution a puzzle); and *involvement* (e.g. getting frustrated vs experiencing satisfaction). Their findings highlight how progress achieved without learning is less satisfying to players, suggesting a close relationship between involvement sand learning [20].

Similar work [19] illustrates how players adopt different strategies in an attempt to overcome breakdowns such as *Experiment* or *Stop & Think*: While this line of research has led to a better understanding of how players learn through overcoming breakdowns and achieving breakthroughs, the mainly qualitative approach adopted is particularly time-consuming to apply and may not be appropriate within a game development context.

Games User Research methods

The research discussed so far has all set out, from an academic perspective, to understand how to best facilitate and assess learning in games, and how to understand and recognize player's problem-solving behaviours. In parallel, much research effort has been devoted to the development of methods for evaluating and improving the player experience, particularly in relation to commercial games. These methods are intended to be used within the game development process and fall broadly under the term of Games User Research (GUR). We argue that, while learning is a key component of the games user experience, existing GUR methods do not take full advantage of recent scholarship on the design and evaluation of learning in games.

User testing is the most common GUR method [5, 21], and involves the direct observation of players as they play. Due to time and resource limitations, user testing often involves relatively small numbers of participants. Observations are typically supplemented with "think-aloud" data (where players are asked to verbalise what they are thinking during play), interviews, questionnaires [21, 33] and/or physiological measurement [30, 41].

Game analytics have also been used to inform GUR, where player behaviour within the game is tracked e.g. metrics such as total playtime, damage dealt per player, etc. [13]. Analytics are used frequently in the context of online games, where the data can be used to continuously refine and redevelop the game. However, as Seif el-Nasr et al. [36] note, there are significant challenges in using analytics for guiding game design; the designer must already know which metrics and which behaviours are important to track, and must invest significant work in interpreting fluctuations in those metrics across players. Thus, while game analytics can be helpful for highlighting points in games where players typically experience difficulties, additional methods are required to examine what exactly it is that players are struggling with and how to overcome those problems.

Similarly, while existing quantitative methods for examining learning progression in educational games [e.g. 2, 16, 32] may prove useful in a commercial GUR context, those methods often require access to large data sets and are arguably less appropriate for earlier prototypes. While there are circumstances where such an analysis has been conducted on a smaller number of players e.g. [17 – where the findings were also triangulated with qualitative data], these approaches still require the application of advanced statistical techniques, such as a hierarchical clustering, and so can be rather complex.

Another GUR method is heuristic evaluation, where games are reviewed according to a checklist of design guidelines [22]. Examples of heuristics include: "The game's interface should be intuitive and easy to use" [39] and "The game is paced to apply pressure without frustrating the players. The difficulty level varies so the players experience greater challenges as they develop mastery" [11]. While heuristic evaluations are quick and cost-effective, they are not based on observed player performance. In addition, they do not normally provide an in-depth focus to on player learning.

An exception to this is the Game Approachability Principles (GAP), a heuristic-based method that most closely matches the aims of the current study. GAP is presented by Desurvire and Wiberg [12] as a set of heuristics for considering learning within tutorials and initial game levels e.g. "Scaffolding failure prevention: player provided with help to meet goals of game". The authors found that a combination of a GAP review with usability testing was particularly effective, where the former helped to uncover approachability issues and the latter was useful for providing more detail on playability issues. However, while the principles were derived from general learning theories, they were not based on empirical studies of how players learn in games. In addition, the focus on applying them to tutorials and initial stages of the game neglects the fact that learning occurs throughout the entire experience of play.

While a combination of GUR methods can provide useful examination of various aspects of the player experience, very few approaches have attempted to evaluate gameplay comprehensively from a learning perspective. This is despite the fact that learning is a significant aspect of engagement and deeper forms of involvement such as ‘flow’ [10]. By overlooking the importance of learning, we are missing an opportunity to gain valuable design insights that could be used to improve the overall player experience.

In summary, the current paper brings together research on how to best design learning in games, with research on how to evaluate learning in games, and, presents it as a useful addition to existing methods for Games User Research. In the next sections, we describe the step-by-step process for evaluating player learning, and present a case study how we applied our method to the puzzle game *Portal*. *Portal* was chosen since it was previously analysed in detail in [25] and is generally considered a very well-designed game.

A METHOD FOR EVALUATING LEARNING DESIGN IN GAMES

In this section, we outline a step-by-step process for identifying problems with the design of learning in games.

Step 1: Chart learning curve of the game

The first step is to create a record of the ideal solution to each puzzle in the game. A list must be created that outlines each action necessary to solve each individual puzzle. In a game development context, designers would be able to supply this information. Next, this list is used to chart the trajectory of complexity across subsequent puzzles (i.e., the pacing or learning curve). In order to chart ‘complexity,’ we use a basic functional definition provide by Linehan et al. [25]. Puzzles that require more actions from the participant are considered more complex than those that require fewer actions. Complexity data is represented on a line chart to chart the trajectory of complexity over progression in the game (see Optimal actions, Fig. 1).

Step 2: Recruit participants, analyse player expertise

In line with existing user testing approaches, small numbers of participants are recruited. The group must also contain sub-sets of players that range in terms of their expertise. Player expertise is measured through a short questionnaire, which establishes familiarity with the game series (if appropriate); frequency of play (how often they play); breadth of experience (the range of genres they play and platforms they use) and gaming history (how long have they been playing games for). Based on results of this questionnaire, players are divided into novice, intermediate and expert categories. This classification is essential, since players with different skill levels are likely to have a range of learning needs. Analysis of player performance will not be useful without knowing their level of expertise. Also, note that the behaviour of novices is probably the best indication of the quality of learning design in a game, since their abilities are derived purely from the observed interactions with the game, and not by information learned in previous play-throughs.

Step 3: Gameplay is recorded in a rich, multi-modal manner

Participants must play the game or game sections that you are interested in. While they do so, it is necessary to record their performance and experience of playing. Many different approaches have been taken in GUR, but for the purposes outlined in this paper it is necessary to record (i) data on all player interactions with the game, allowing for quantitative analysis of player performance, and (ii) data on player experience in terms of how they solve the problems presented by the game e.g. concurrent or retrospective think aloud to allow for more in-depth qualitative analysis.

Step 4: Analyse participant performance

Whether player performance was recorded automatically through data logs (ideal) or manually through video recordings (more time intensive), performance data for each participant is extracted and represented on the same graphs from step 1. In this way, the performance of each player can be compared with the optimum solution. In addition, statistical means should be created for each player type (novice, intermediate, expert) and represented on the graphs.

Step 5: Identify game sections where problems have arisen

A well-designed game builds competence in players gradually and methodically, so that players are never presented with challenges that are significantly beyond their ability. Thus, a competent player is one who, when presented with a puzzle, regularly comes to the solution efficiently. Conversely, a player lacking in competence, who hasn’t fluently learned the appropriate skills, will regularly face breakdowns in their gameplay and demonstrate problem-solving behaviour that resembles trial and error. A well-designed game should generally elicit competent behaviour from participants, as this is evidence of the player learning and applying skills effectively. The graphs that emerge from step 4 will identify sections of the game (or individual puzzles) where there is a consistent and large divergence between the ideal solution and the performance of players. These are the sections that must be investigated further.

Step 6: Conduct qualitative analysis of those problems

A detailed qualitative analysis of data relating to the problem sections of the game is carried out. The process follows a ‘critical incidents’ methodology, gathering data from multiple sources to understand the player experience relative to that one event. To facilitate coding, a table is created for each player who experienced difficulties in the problem section. The table lists the actions the player carried out (observed from the video) and any relevant think aloud and post-play interview quotes. The analysis is deductive, utilising categories from previous research, where separate columns are used to code *breakdowns* [20] and *strategies* [19]. Breakdowns are categorised as relating to *action* (when a player fails to execute an action successfully), *understanding* (when the player is confused what to do next or is suffering from a misunderstanding and *involvement* (when they appear bored or frustrated). Breakdowns are classified as *‘major’* when recurring across multiple players or when they significantly obstacle player progress (e.g. the player ‘dies’ or is stuck for an extended time period). All other instances are *‘minor’* breakdowns. Strategies are coded

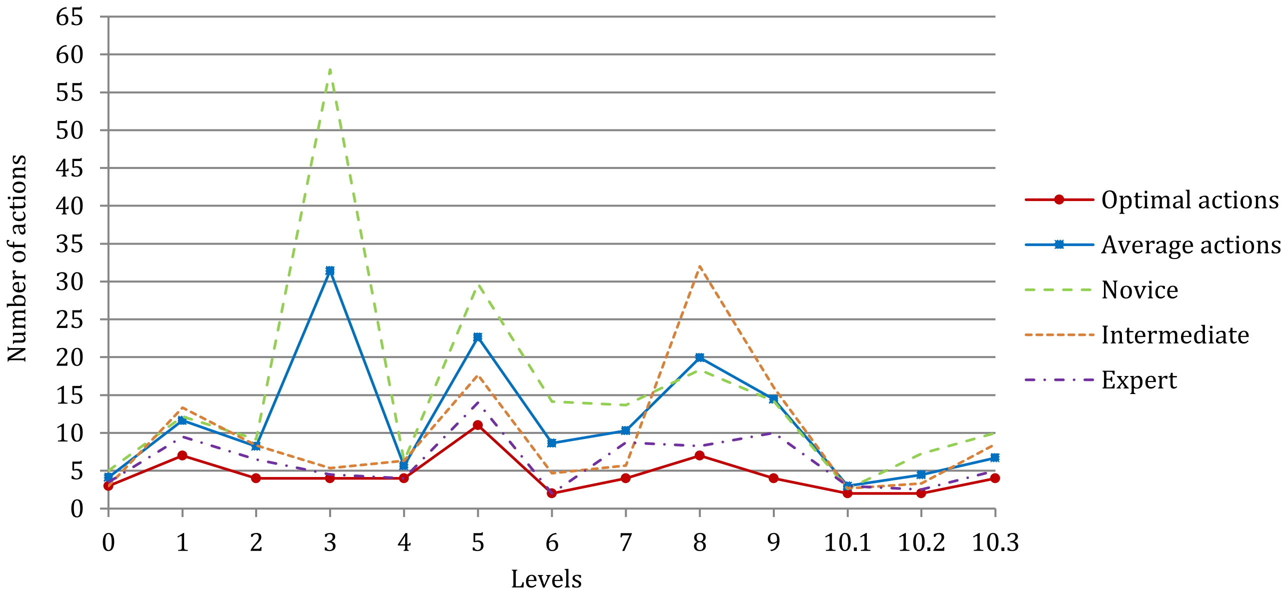

Figure 1: Action curves for (1) Optimal (minimum number of actions); (2) average actions across groups; (3) average actions for novices; (4) average actions for intermediates; (5) average actions for experts.

as *Trial & Error (*where a player carries out an action to see what, if anything will happen); *Experiment* (where the player has an informal theory about what will happen when they try something), *Stop & Think* (gameplay is suspended to reflect or consult external resources); *Take the Hint* (when the player decides to follow guidance provided within the game); and *Repetition* (when the same actions are practiced or repeated by the player several times).Through examining sets of tables related to each problem section within a the game, a deeper understanding can be developed of where difficulties are occurring and why.

Step 7: Develop design recommendations

After considering the causes of major breakdowns as part of the qualitative analysis, design recommendations can be distilled and discussed with the developers as suggestions for improving the learning design of the game.

CASE STUDY

Participants were asked to play a single-player puzzle videogame, *Portal* [42]. The game employs a 3D first-person perspective, where players have to overcome a variety of puzzles in self-contained 'test chambers' using portals to create pathways e.g. between the player and an otherwise inaccessible platform. The levels often contain visual hints on the ground and walls, while the voice of a robotic AI called GLaDOS (Genetic Lifeform and Disk Operating System), initially guides the player (though she becomes increasingly malicious as the game progresses). The game gradually introduces mechanics to the player, progressing from initial levels, where the test chamber setup controls the portals, into more complex levels, where the player is then able to shoot one portal type (blue), and eventually allowed to shoot both types (blue and orange). Simple instructions sheets regarding the controls were provided to players.

Step 1: Chart learning curve of the game

The optimum solutions for *Portal* were taken from Linehan et al. [25] where, a list has been published of all actions necessary to solve each puzzle in the game. These data were used to chart the optimal learning curve (see Fig. 1).

Step 2: Recruit participants, analyse player expertise

Fourteen participants were recruited from a university participant pool (5 female, 9 male, average age 25.1 years). All participants reported playing games at least once a week, and were recruited on the basis that they were familiar with games involving a first-person perspective (mainly FPS games) and playing with console controllers. Volunteers who played infrequently and solely on mobiles or tablets were excluded to minimise action breakdowns due to a lack of familiarity with controls. Based on their prior gaming experience and familiarity with the *Portal* series [42, 43], as evidenced by filling in a short questionnaire, participants were classified as 'Experts' if they had completed *Portal* before (N = 4); 'Intermediates' if the player had some exposure to *Portal*, or had completed *Portal 2* only (N = 3); 'Novices' if they had never or barely played either game (N = 7). Players are referred to by expertise and number e.g. PI08 is Player Intermediate 08.

Step 3: Gameplay is recorded in a rich, multi-modal manner

The study was setup in a professional UX laboratory, like a living room, with a TV and sofa. The game was played on an Xbox 360 console, with a wired controller. A video camera captured a recording of the participant, and the console connected to a recording facility to capture game footage. Microphones were used to record comments, and Media Express software for capturing all inputs.

Participants played the game for 40 minutes and asked to explain what they were doing and thinking whilst playing. A researcher was present to prompt them if they became silent. After the session, participants were asked to how much they agreed with the statement "I have enjoyed the game". The majority were positive, with 9/14 selecting 'strongly agree' and 3/14 selecting 'agree' and only 2/14 selecting 'neutral'. A short interview (10-15 minutes) then took place after the session where players were asked to review instances where they appeared to be struggling and asked to discuss these in more detail.

Step 4: Analyse participant performance

Figure 1 includes the optimal learning curve for *Portal* [25], showing the minimum number of actions required for each level. In addition, the figure displays the overall player performance curve (the total average actions per player who completed each level); and the player performance curves for each group (average actions for Novices, Intermediates and Experts). We focus on the first 10 levels of the game as less than two players from each group had progressed beyond this point within the 40-minute session.

The optimal curve indicates that *Portal* is a relatively well-designed game in that player behaviour appears to follow the general pattern of the optimal curve. However, it is clear there are some large discrepancies between optimal and average performance, particularly in relation to group expertise. While Experts tend to mirror the optimal action curves very closely (which is not surprising in this context as they were all very familiar with the *Portal* series), the discrepancies between the optimal and average curves indicate that Novices are struggling with Levels 3, 5 and 6 while Intermediates had difficulties in Level 8.

Step 5: Identify game sections where problems have arisen

We suggest that large discrepancies between the curves can serve as evidence of problems with how the game was designed to support learning. Levels 3, 5, 6 and 8 (see Table 1) showed the largest differences between optimal and average performance for particular groups so they were selected for further analysis.

	Total Average		Novice		Intermediate		Expert	
	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N
L.3 [4]	31.4 (36.9)	14	58 (36)	7	5.33 (2.3)	3	4.5 (1)	4
L.5 [11]	22.6 (19)	14	29.7 (25.6)	7	12.7 (2.1)	3	14 (4.1)	4
L.6 [2]	8.64 (9)	14	14.1 (9.8)	7	4.7 (4.6)	3	2 (0)	4
L.8 [7]	19.9 (13.8)	12	22 (6.8)	5	32 (21.3)	3	8.3 (1)	4

Table 1: Total average and group actions for Levels 3, 5, 6, 8 with optimal actions in []

Step 6: conduct qualitative analysis of those problems

The quantitative analysis made clear that there were significant issues that required investigation in levels 3, 5, 6 and 8. Video, think aloud and interview data were analysed in terms of breakdowns that occurred and the strategies players adopted (as outlined above).

Strategies players employ when experiencing breakdowns

The pattern of strategies employed by players experiencing major breakdowns was relatively consistent and serve as a clear indicator of when they were struggling. Gameplay primarily consisted of *Trial & Error* and *Experiment* strategies as players tried to understand the game world and rules. When experiencing difficulty, players often resorted to a combination of *Trial & Error* actions e.g. firing portals at different objects in the room, and *Repetition*, when they repeated the same action several times. In addition, the think-aloud indicated when they were forming incorrect hypotheses during *Experiment*, such as thinking that creating a blue portal in a different part of the wall would change the location of the exit. Occasionally, players would *Stop & Think* where they would become inactive in the game and reflect on what they were doing or looked at menus and external resources for further hints. If a player didn't initially *Take the Hint* (in the form of visual in-game clues) but then went back to inspect them, this was seen as further evidence that they were struggling. In the subsections below, each level is introduced before examining the nature of the breakdowns experienced.

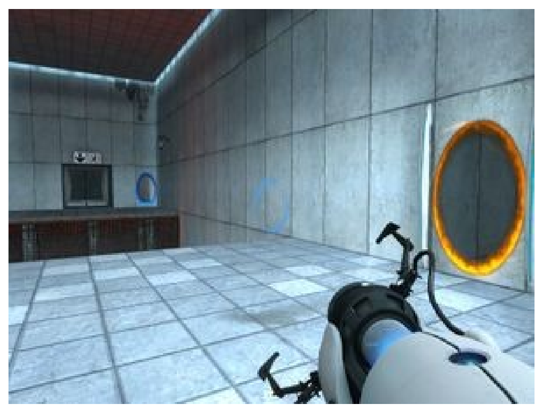

Figure 2: Level 3

Level 3: Entering and exiting different coloured portals

At the entrance of this L-shaped level, a pit separates the player from a middle floor with a fixed orange portal at the end. Around the corner from the middle floor another pit that separates this section from the exit on a higher floor. The optimal solution requires 4 steps: firing a blue portal close to the entrance; walking through it to exit out the orange portal on the middle floor (a); firing a blue portal near the distant exit (b); and walking into the orange portal to exit from the blue portal on the other side (Fig. 2). Novice players struggled with this level, taking an average of 58 actions to complete it (SD = 36), which is 14.5 times the minimum number of actions required (see Table 1).

In the previous level, a new mechanic was introduced in the form of a portal gun that allowed players to create blue portals. While all players were able to use the gun to reach the middle floor, once they were there, the majority of novices started to iterate the same set of actions. For example, PN07 reached the middle floor, commented “No idea where I am”, then turned, looked through orange portal, and fired a blue portal next to it. Here the player starts to repeat the sequence of ‘fire blue’+’enter blue’, soon realising that those actions always led to “the same place”. Similarly, PN11 repeatedly fires and enters a blue portal, noting “wherever I make a portal it still leads to the same place”. Some players also tried shooting portals in each of the lateral walls, thinking that changing the entrance of the portal would lead to a different exit e.g. PN05 explains he “tried open portals on all the walls, and it just took me around in a circle”.

PN01 and PN05 experienced further confusion when they thought they had encountered another character in the game. However, this is actually their own avatar, which can be viewed due to the way the portals are lined up e.g. PN05: “I can follow her. I am sure she can get me there”, entered and exited the portals several times: “I am following this woman, but she is going in circles”, before eventually realising “Oh no, that’s my reflection, she’s me”. Another indication of difficulty in this level was that several of them suspended gameplay to look up the controls of the game to see if they were neglecting possible actions.

Level 3: Summary of main issues

Novice players seemed to be experiencing a major understanding breakdown around how the portals work. While these players frequently entered the blue portals they created, they only occasionally entered the fixed orange portal – not seeming to realise it could be both exit and entrance e.g. PN01 “it’s blue when I enter, orange when I get out”. Though all players eventually managed to complete the level, for some this was accidental as they did not understand how they reached the exit. For instance, PN07 wonders aloud “I went the right way, it seems, I think … but I don’t know how I did it” while PN04 utters a surprised “Oh. How did that happen? … this is completely random”. The breakdowns experienced and the fact some players could progress without gaining this knowledge suggests that the learning design of the level could be improved, at least for novices.

Level 5: More trouble with portals

Level 5 (Fig. 3) presents the player with a large room which contains a door activated by two buttons (each of which require a cube), two raised platforms facing each other and a pit which contains another cube. The door gives access to a second room and closes after the player walks past it. In the second room, a new orange portal is visible through a glass ceiling of the room above. The optimal solution requires 11 steps: firing a blue portal under the first cube in the pit; entering the blue portal within the pit; picking up the cube on the platform with the orange portal; placing the cube on one of the buttons; entering the blue portal again to go back up to the platform; firing a new blue portal on the platform opposite; entering the orange portal; picking up the second cube; placing the second cube on the remaining button; going through the door into the second room and shooting a new blue portal into the wall; entering the portal to exit out of the orange portal visible through the glass. Novices players were also seen to struggle with this level, taking an average of 29.7 actions to complete it (SD = 25.6), which is 2.7 times the minimum number of actions required (Table 1).

Figure 3: Level 5

The main reason for the large action count within this level is due to the fact some of the players were still suffering from a major understanding breakdown concerning how the portals worked. For instance, PN05 explains “I thought maybe if I try this wall, maybe a portal will lead me to here [platform with the cube]” suggesting that he believes the orange portal is only for exiting and that it can perhaps be controlled by changing where the blue portal is fired. After accidentally backing in to the orange portal, the player is able to grab the second cube “Ok, I got it” but again, does not seem to be able to explain why.

PN04 follows a similar strategy to one observed in level 3, shooting at different panels to “see whether there is a hidden area in the wall that will lead to that place, the one where the other cube is”. After some time, when the player has created a blue portal on the ledge she wants to get to, she turns around and realises she can go through the orange portal exclaiming “Oh, this is how it works!”. However, when asked to elaborate on what happened she seems a little unsure: “I don’t know how to explain it but its’ like you need to have one portal in one place and another one pointing to that so it creates a doorway and that’s how it works”. PN07 also struggled, though after he reaches the platform via the orange portal has a breakthrough: “Doh! I had to do the double portal thing again”, remembering that you can go “back and forth through the fixed portal”.

Level 5: Summary of main issues

While this level was more complex than Level 3, as indicated by the higher optimal action count, these examples show how a major breakdown in understanding that begins in an earlier level will influence later gameplay.

Level 6: Introducing High Energy Pellets

In Level 6 players encounter a new mechanic in the form of the High Energy Pellet (HEP) emitter which produces a fast-moving energy ball which the players must direct using the portal gun. Energy pellets eventually dissipate, and when they do the receiver emits a new one. The level composes of a HEP emitter on the ceiling, an inactive orange portal opposite to it, a receiver to the right of the inactive orange portal and a vertical scaffold activated by a HEP hitting the receiver (Fig. 4). The optimal solution requires 2 steps: fire a blue portal on the ceiling above the receiver (to guide the HEP to it); walk onto the scaffold which will then rise to the exit. Most of the novice players struggled with this level, taking an average 14.1 actions to complete it (SD = 9.8), 7.1 times the minimum actions required (Table 1).

Figure 4: Level 6

When entering Level 6, players are introduced to the HEP by GLaDOS who states, “While safety is one of many Enrichment Center goals, the Aperture Science High Energy Pellet, seen to the left of the chamber, can and has caused permanent disabilities, such as vaporization”. However, while some participants realised straight away that the HEP would be dangerous (e.g. PN11: “ok, so I can’t touch that thing”) others ended up dying, sometimes more than once, by walking right up to the pellet and coming into contact with it. For instance, PN01 struggled with the GLaDOS voiceover: “I didn’t understand anything about the voice that was talking to me” as did PN07: “Not sure what they are referring to here”. PN12 seemed to ignore the voiceover altogether and mistakenly assumed “I thought I had to collect the pellet, because I thought it was similar to the blue trigger” [referring to picking up the portal gun in a previous level].

Apart from experiencing an understanding breakdown regarding the nature of the HEP, there was further confusion about how to solve the puzzle. Most players quickly realised they could not create portals on black walls, but then resorted to shooting at both the emitter and receiver objects and the closed orange portal (e.g. PN01, PN05, PN11, PN12, PI14). In addition, PN05 initially thought that portals could “only open on the ground” while PN01 only realises it’s possible to shoot portals in the ceiling when trying to shoot the emitter and accidentally creating a portal nearby: “I opened a door… Ah I got it”.

Level 6: Summary of main issues

While some players did look at the hints available in the game, including those at the start of the level and on the ground, none of the players who struggled seemed to notice the red light emanating from the emitter that pointed up towards the ceiling, which acts as a cue of for positioning a blue portal. Instead, they adopted strategies such as PN12 who walked on top of the receiver and looked directly up to fire a portal directly above. The breakdowns experienced in this level suggest the information provided by GLaDOS was not clear to some players and that some of the cues provided were not salient enough.

Figure 5: Level 8

Level 8: Issues with perspective

In Level 8, the Test Chamber floor is covered by lethal water and presents an inactive portal on a platform on the left, and a horizontal scaffold on the right. On the left, there is a HEP emitter and a fixed orange portal on a platform, and on the right a receptacle on the wall, which will activate the immobile scaffold (Fig. 5). The level requires a minimum of 7 actions: shoot blue portal opposite the HEP dispenser; after the HEP passes through, create a new blue portal opposite the receptacle to activate the scaffold; once activated, shoot a blue portal in the wall nearby the entrance; travel through the portal to exit out of the fixed orange portal; create a blue portal above the scaffold and wait for it to appear; go through the orange portal; ride scaffold and proceed to exit. Intermediate players had difficulties with this this level, needing an average of 32 actions to complete it (SD = 21.3), which is 4.6 times the optimal number of actions required. Some novices also struggled here, requiring an average of 22 actions (SD = 6.8) to complete the level (Table 1).

The main breakdown in this level (experienced by players PI08, PI14 and PN11) originates from the central perspective of the chamber, which provides the illusion that the HEP receiver is located opposite the orange portal. P108 was able to quickly realise a solution to the problem but decided to look through the orange portal to figure out exactly far away the receptacle is on the other side. Unfortunately, this strategy leads to her dying twice and spending a large amount of time trying to count how many panels away the receptacle is e.g. “one, two, three, four…” before she successfully locates a portal opposite it. PI14 describes this level as “horrible” noting “it was the angle that was the problem”; their reaction to the level could be seen as an instance of a potential involvement breakdown. PN11 noted the same difficulty though was able to overcome it stopping for a moment and taking a different perspective: "I feel I am looking at this by the wrong angle {moves}. Oh! The machine is here! That makes a lot more sense".

Level 8: Summary of main issues

As in Level 6, a subtle red flare is projected by the receptacle onto the opposite wall that could have helped with portal positioning. However, the cue was again missed and the initial perspective the players encounter made this a particularly frustrating experience for some as it provided an obstacle to realising the solution to the puzzle.

Step 7: Develop design recommendations

Through considering the causes of breakdowns as part of the qualitative analyses, there are three main areas where the design could be improved to support learning:

1. Understanding of how portals work (Levels 3 & 5)

The Developer Commentaries [44] show that a design decision was made to force players to enter the orange portal in Level 3 by making it a fixed entrance. Despite this attempt, the fact that players could progress in this level without fully understanding how the portals operated caused a major breakdown. This is evidenced through observations of exactly the same errors being made in Level 5. This breakdown seems to stem from Level 0, where players didn't understand they could see themselves in portals and continues through subsequent levels. Potential solutions could involve locating a mirror in the room where the player starts so they see their reflection before they see themselves in a portal, providing them with a gun that can shoot both blue and orange portals in Level 2, or changing the map in Level 3 from a "L-shape" to an "I-shape" to reduce the complexity of the level.

2. Making cues more salient (Levels 6 & 8)

There may be some scope to make GLaDOS's voiceover clearer for the introduction to Level 6, but, in general, players learned relatively quickly that the HEPs were dangerous. The main issue in this level related to the fact that many did not attempt to shoot a portal in the ceiling until they had exhausted other options. Though hints were provided, these were perhaps a little too abstract, while the red light (provided to cue players to shoot above the receiver) was unfortunately not obvious. The same is true for Level 8, where again the red light pointing from the receiver to the opposite wall was rather faded (as Fig. 8 indicates). These issues suggest that the designers may need to alter the saturation and luminosity of the red light to make it more salient on console versions where television resolution could reduce the visibility of the cue and prevent players from achieving an understanding breakthrough that could help them in later levels. Unfortunately, these difficulties could be beyond developer control as in the case of playing the Xbox version of *Portal*, where the "quality of the graphics has been downsized to meet the restricted game size of Microsoft's XBLA service" [31]. Nevertheless, the use of red colour on grey walls creates an accessibility barrier for certain types of colour blindness and colours could be chosen with this in mind. Another option would be to provide players with substitute palettes on the whole game.

3. The issue of perspective (Level 8)

Level 8 was designed to teach the redirection technique [44] and players appeared at ease in grasping the concept. However, in addition to the fact that the red-light cue was not very visible, the perspective players were provided with when walking into the level mislead them to think that the receiver was located closer to the orange portal that it was. The design of Level 8 could be altered by moving the entrance point to force player to gain a different perspective when they enter the chamber, thus increasing their chances of figuring out the correct solution to the puzzle.

DISCUSSION

This paper represents an effort to devise a comprehensive evidence-based method for identifying problems with how learning is structured within commercial games, and for offering evidence-based solutions to these problems. The approach is based on a mixing of two strands of contemporary research on learning in games; one quantitative and behavioural, the other qualitative and experiential. The method consists of seven steps:

1. Chart learning curve of the game
2. Recruit participants, analyse player expertise.
3. Gameplay is recorded in a rich, multi-modal manner.
4. Analyse participant performance.
5. Identify game sections where problems have arisen.
6. Conduct qualitative analysis of those problems.
7. Develop design recommendations.

The method presented directly builds on previous work by Linehan et al. [25] and Iacovides et al. [19] by combining a learning curve analysis with work on gameplay breakdowns into a comprehensive method for evaluating how games teach players game-specific skills and information. While existing GUR methods recognise the importance of learning design to the player experience, and include some evaluation criteria that relate to learning e.g. [12], we suggest that the method we propose is particularly capable of detecting subtle problems with player understanding (across the whole experience of playing a game) and providing clear solutions to those problems. The method also fits seamlessly the type of small-scale user test frequently undertaken in early development phases, without requiring the large-scale data acquisition and advanced statistical or visualisation techniques that are involved in analytics approaches [e.g. 2, 16, 17, 32]. The approach could be used throughout the development process – from being applied to evaluate learning difficulties around key game mechanics in low fidelity prototypes to evaluating the impact of specific design elements within more developed versions.

Linehan et al. [25] suggest that *Portal* is a well-designed game, as it strictly follows a set of instructional design principles based firmly in the behavioural education literature (see [26]). Our evaluation of *Portal* provides new empirical evidence to support this claim, as players generally solved puzzles in an efficient manner, readily demonstrating the appropriate application of skills learned within the game. Furthermore, participants generally reported enjoying playing it. However, our analysis did also reveal a small number of issues, where large deviations between curves indicated areas where the game could be improved upon, particularly for novice players.

Part of the process we outline was to recruit participants who varied in terms of their expertise. This helped to ensure that we took into account the needs of different types of players. Indeed, it could be argued that the behaviour of novice players is the most informative in terms of assessing the quality of learning design in a game. However, it is important to also recruit more experienced players to understand whether performance improves on repeated play-throughs. Finding that experts perform only marginally better than novices would indicate significant issues with how the learning process is supported.

Furthermore, expertise can be hard to classify [19, 24]. We settled on a combination of questions to assess participants that related to not just how often they play, but what genres and platforms and for how long they had played. Nevertheless, some who were novices managed to complete more of the game within the play sessions than intermediates i.e. they showed more competence at solving the puzzles despite having less general experience. Researchers and practitioners need to be clear about the criteria they use to classify their participants while the area would benefit from clear guidelines about how to do so.

Limitations

This method is presented as useful in the context of puzzle games. While we only evaluated *Portal*, the method should be applicable to other games that have a set of features that allow them to benefit from a structured analysis. For example, puzzles with 'optimum' solutions which are can be compared to player behaviour. This may not be the case with simulation-based games that have multiple outcomes, or games that reward player creativity. While not impossible, it would be more difficult to plot the curves in a more 'open world' that gives agency to players to choose the order in which they encounter puzzles and challenges.

Additionally, due to the inclusion of rich qualitative data, the method is only appropriate for using with a relatively small number participants. While that may increase the chances of the method falling prey to idiosyncrasies in play testing populations, the same can be said of most small-scale user testing scenarios. An extreme outlier could be removed if a particular participant was seen to skew the results but the general aim of similar GUR approaches is to highlight and reduce potential player experience issues rather than produce statistical generalizations about frequencies and populations.

Future work

Where a large divergence between the optimum curve and that achieved by players, we suggest that difficulty should be reduced to avoid the risk of players becoming frustrated and deciding give up on the game. Conversely, in puzzles where player performance followed the curve very closely, gameplay is arguably too easy and may become boring. Ideally then, there should be some distance between the optimal and player curves. The crucial question for further research is, how much distance should there be between the optimal and average observed curves? This question has repeatedly obsessed game design researchers, as evidenced through large quantities of theoretical and empirical work on concepts such as 'flow' 'immersion' 'appropriate challenge', with a recent focus on using learning analytics. We are yet to see a unifying theory of learning design in games, but the current paper does provide a method through which those questions can be asked and answered of specific games.

In addition, the data collected can provide some insight into player engagement. Though there was little evidence of major involvement breakdowns within this study, observational, think aloud and interview data could indicate when players are becoming frustrated or bored with a section of the game [e.g. see 19]. A focus on real-time involvement would be particularly important to carry out when applying the method as part of the game development process.

While have presented our work as a case study to illustrate how our approach can be implemented in practice, further research is also required to examine how applicable it is to other games and genres. Furthermore, it would be useful to apply the method within the context of game development to validate the process and to assess feasibility in practice. Other studies could also carry out a comparative analysis with other GUR methods to explore how effective the approach is in terms of the quantity and quality of design issues it is able to uncover.

CONCLUSION

We present a mixed method approach to uncovering problems with how player learning is supported in games. The method is based in recent academic research on designing and evaluating learning in games. A case study is presented, where the method is applied to *Portal*. Issues are identified with a small number of the puzzles in that game, and explored through an in-depth qualitative analysis. While more work is required to examine the feasibility of using this method within an ongoing game development process, this paper represents an initial attempt at a comprehensive evidence-based method for identifying problems with how learning is structured within games, and for offering clear evidence-based solutions to these problems.

ACKNOWLEDGMENTS

We thank all the volunteers who took part in both studies. This research was supported by the EPSRC funded CHI+MED project (EP/G059063/1).

REFERENCES

1. Justin T. Alexander, John Sear and Andreas Oikonomou. 2013. An investigation of the effects of game difficulty on player enjoyment. *Entertainment*

Computing, 4, 53–62. http://dx.doi.org/10.1016/j.entcom.2012.09.001

2. Erik Andersen, Yun-en Liu, Ethan Apter, François Boucher-genesse, and Zoran Popovi. 2010. Gameplay Analysis through State Projection. Proceedings of the 5th International Conference on the Foundations of Digital Games - FDG '10, ACM Press, 1–8. http://dx.doi.org/10.1145/1822348.1822349
3. Erik Andersen, Eleanor O'Rourke, Yun-En Liu, Rich Snider, Jeff Lowdermilk,David Truong, Seth Cooper, and Zoran Popovic. 2011. The impact of tutorials ongames of varying complexity. In *Proc. CHI '12*, ACM Press, 59-68.
4. Fran. C. Blumberg, Sheryl F. Rosenthal and John D. Randall. 2008. Impasse-driven learning in the context of video games. *Computers in Human Behavior*, 24(4), 1530-1541. http://dx.doi.org/10.1016/j.chb.2007.05.010
5. Emily Brown. 2010. The life and tools of a games designer. In *Evaluating User Experience in Games: Concepts and Methods*, Regina Bernhaupt (ed.). Springer London, 73-87. http://dx.doi.org/10.1007/978-1-84882-963-3_5
6. Paul Cairns, Anna L. Cox, and Imran Nordin, I. 2014. Immersion in digital games: a review of gaming experience. In *Research in Handbook of Digital Games*, Marios C. Angelides and Haryy Agius (eds.). IEEE/John Wiley and Sons, 339-361. http://dx.doi.org/10.1002/9781118796443.ch12
7. Marcus Carter, John Downs, Bjorn Nansen, Mitchell Harrop and Martin Gibbs. 2014. Paradigms of games research in HCI: a review of 10 years of research at CHI. In *Proceedings of the 1st ACM Conference Computer-Human Interaction in Play* (CHI PLAY '14), 27-36. http://dx.doi.org/10.1145/2658537.2658708
8. Alessandro Canossa. 2009. *Play-Persona: Modeling Player Behavior in Computer Games.* Copenhagen, Danish Design School Press.
9. Daniel Cook. 2007. The Chemistry of Game Design. *Gamasutra*. Retrieved April 14, 2017 from: http://www.gamasutra.com/view/feature/129948/the_chemistry_of_game_design.php?print=1
10. Mihaly Csikszentmihalyi. 1990. Flow: The psychology of optimal performance. New York: Harper and Row.
11. Heather Desurvire and Charlotte Wiberg. 2009. Game Usability Heuristics (PLAY) for Evaluating and Designing Better Games: The Next Iteration. In *Proceedings of the 3d International Conference on Online Communities and Social Computing: Held as Part of HCI International 2009*, OCSC '09, 557–566. http://dx.doi.org/10.1007/978-3-642-02774-1_60
12. Heather Desurvire and Charlotte Wiberg 2010. User Experience Design for Inexperienced Gamers: GAP - Game Approachability Principles. In *Evaluating User Experience in Games: Concepts and Methods*, Regina Bernhaupt (ed.). Springer London, 131-147. http://dx.doi.org/10.1007/978-1-84882-963-3_8
13. Anders Drachen, Magy Seif El-Nasr, and Alessandro Canossa. 2013. Game Analytics – The Basics. In *Game analytics: Maximizing the value of player data*, Magy Seif El-Nasr, Anders Drachen and Alessandro Canossa (eds.). London, UK: Springer-Verlag, 13–40. http://dx.doi.org/10.1007/978-1-4471-4769-5_2
14. James P. Gee. What Video Games Have to Teach us about Literacy and Learning. 2004. New York: Palgrave Macmillan.
15. James P. Gee. 2005. Learning by Design: Good video games as learning machines. *E–Learning, 2*, 5-16. http://dx.doi.org/10.2304/elea.2005.2.1.5
16. Erik Harpstead and Vincent Aleven. 2015. Using Empirical Learning Curve Analysis to Inform Design in an Educational Game. In *Proceedings of the 2015 ACM on Computer-Human Interaction in Play* (CHI PLAY '15), 197-207. http://dx.doi.org/10.1145/2793107.2793128
17. Britton Horn, Amy K. Hoover, Jackie Barnes, Yetunde Folajimi, Gillian Smith and Casper Harteveld. 2016 Opening the Black Box of Play: Strategy Analysis of an Educational Game. In *Proceedings of the 1st ACM Conference Computer-Human Interaction in Play* (CHI PLAY '16), 142-153. http://dx.doi.org/10.1145/2967934.2968109
18. Ioanna Iacovides, James C. Aczel, Eileen Scanlon, Will I.S Woods. 2011. What can breakdowns and breakthroughs tell us about learning and involvement experienced during game-play? In *Proceedings of the 5th European Conference of Game-based Learning*, 275-281, Academic Publishing Limited.
19. Ioanna Iacovides, Anna L. Cox, Ara Avakian and Thomas Knoll. 2014. Player strategies: Achieving breakthroughs and progressing in single-player and cooperative games. In *Proceedings of the 1st ACM Conference Computer-Human Interaction in Play* (CHI PLAY '14), 131-140. http://dx.doi.org/10.1145/2658537.2658697
20. Ioanna Iacovides, Anna L. Cox, Patrick McAndrew, James C. Aczel and Eileen Scanlon. 2015. Game-play breakdowns and breakthroughs: Exploring the relationship between action, understanding and involvement. *Human Computer Interaction*, 30 (3-4), 202-231. http://dx.doi.org/10.1080/07370024.2014.987347.
21. Katherine Isbister and Noah Schaffer. 2008. *Game Usability: Advancing the Player Experience*. Morgan Kaufman.
22. Christina Koeffel, Wolfgang Hochleitner, Jakob Leitner, Michael Haller, Arjan Geven and Manfred

Tscheligi. 2010. Using heuristics to evaluate the overall user experience of video games and advanced interaction games. In *Evaluating User Experience in Games: Concepts and Methods*, Regina Bernhaupt (ed.). Springer London, 233-256. http://dx.doi.org/10.1007/978-1-84882-963-3_13

23. Raph Koster. 2005. *Theory of fun for game design*. Scottsdale, AZ: Paraglyph Press.
24. Andrew J. Latham, Lucy L. M. Patston and Lynette J. Tippett. 2013. Just how expert are "expert" video-game players? Assessing the experience and expertise of video-game players across "action" video-game genres. *Frontiers in Psychology*, 4, Article 941. http://dx.doi.org/10.3389/fpsyg.2013.00941
25. Conor Linehan, George Bellord, Ben Kirman, Zachary H. Morford, Bryan Roche. 2014. Learning curves: analysing pace and challenge in four successful puzzle games. In *Proceedings of the 1st ACM Conference Computer-Human Interaction in Play* (CHI PLAY '14), 27-36, 181-190. http://dx.doi.org/10.1145/2658537.2658695
26. Conor Linehan, Ben Kirman, Shaun Lawson, & Gail. C. Chan. 2011. Practical, appropriate, empirically-validated guidelines for designing educational games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, 1979-1988. http://dx.doi.org/978-1-4503-0267-8/11/05
27. Geoffrey R. Loftus and Elizabeth F. Loftus. 1983. *Mind at Play: The Psychology of Video Games*. New York, NY: Basic Books.
28. David B. Malouf. 1987. The effect of instructional computer games on continuing student motivation. *Journal of Special Education, 21*, 27-38. http://dx.doi.org/10.1109/1177/002246698802100406
29. Merrilea J. Mayo. 2007. Games for science and engineering education. *Communications of the ACM, 50* , 31-35. http://dx.doi.org/10.1145/1272516.1272536
30. Pejman Mirza-Babaei and Lennart Nacke. 2014. Introducing the biometric storyboards tool for games user research. *Games Media Entertainment* (GEM), IEEE, 1-7. http://dx.doi.org/10.1109/GEM.2014.7048098
31. MobyGames (nd). Portal: Still Alive (Xbox). Retrieved April 14 2017 from: http://www.mobygames.com/game/portal-still-alive
32. Elizabeth Owen, Dennis Ramirez, Allison Salmon, and Richard Halverson. 2014. Capturing Learner Trajectories in Educational Games through ADAGE (Assessment Data Aggregator for Game Environments): A Click-Stream Data Framework for Assessment of Learning in Play. *American Educational Research Association Annual Meeting*: 1–7.
33. Randy J. Pagulayan, Kevin Keeker, Dennis Wixon, Ramon L. Romero and Thomas Fuller. 2002. User-centered design in games. In *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications*¸ Julia A. Jacko & Andrew Sears (eds.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc, 883-906).
34. Caroline Pelletier and Martin Oliver. 2006. Learning to play in digital games. *Learning, media and technology*, 31, 329-342. http://dx.doi.org/10.1080/17439880601021942
35. William Ryan and Martin A. Siegel. 2009. Evaluating interactive entertainment using breakdown: Understanding embodied learning in video games. In *Proceedings of Digital Games Research Association*, (DIGRA '09).
36. Magy Seif El-Nasr, Heather Desurvire, Bardia Aghabeigi, Anders Drachen. 2013. Game analytics for game user research, Part 1: A workshop review and case study, *Computer Graphics and Applications*, IEEE 33 (2), 6-11. http://dx.doi.org/10.1109/MCG.2013.26
37. Mike Sharples. 2009. Methods for evaluating mobile learning. In *Researching Mobile Learning: Frameworks, Tools and Research designs*, Giasemi Vavoula, Norbert Pachler, Agnes Kukulska-Hulme (eds.). Oxford: Peter Lang Verlag, 17–39.
38. William Swartout, and Michal Van Lent. 2003. Making a game of system design. Communications of the ACM, 46, 32-39. http://dx.doi.org/10.1145/792704.792727
39. Penelope Sweetser, Daniel M.Johnson and Peta Wyeth. 2012. Revisiting the GameFlow model with detailed heuristics. *Journal of Creative Technologies*, 3.
40. Penelope Sweetser and Peta Wyeth. 2005. GameFlow: A Model for Evaluating Player Enjoyment in Games. *ACM Computers in Entertainment*, 3(3), Article 3A. http://dx.doi.org/10.1145/1077246.1077253
41. Chek Tien Tan, Tuck Wah Leong and Songjia Shenan. 2014. Combining think-aloud and physiological data to understand video game experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 381-390. http://dx.doi.org/10.1145/2556288.2557326
42. Valve Corporation. 2007. *Portal.* Valve Corporation.
43. Valve Corporation. 2011. *Portal 2.* Valve Corporation.
44. Valve Corporation. 2007. *Portal: Developers Commentaries.* Retrieved April 14 2017 from: https://www.youtube.com/watch?v=2TBIfsPGeQ